

Dr. Mark Davis  
President  
Unicode Consortium

L2/07-062

February 5, 2007

Dear Sir,  
Subject: Malayalam Encoding in Unicode

Unicode encoding of Indic scripts is derived from ISCII. ISCII is the basis of a large number of software, from word processors and sorting engines, to morphological analysers and spell checkers. So far, no one has claimed that there is any problem for Malayalam in ISCII for the purpose of computing. ISCII was developed by a team of several national level organizations including CDAC.

It is also to be noted that it is except for one person in the Malayalam Language Technology group at CDAC, no problems have been found by other members or allied groups in CDAC for the representation of chillus using ZWJ, or their equivalent representation in ISCII in applications produced by them.

Dear UTC members, it is only a few people who revolve around the pillars of power, argue for the chillu encoding. At a point in the past, they made some mistakes which were then heavily criticized; now it has become an issue of prestige. The ministers and political advisors are not trained linguists or computer experts, and so are not at all in a position to evaluate these claims, and we cannot blame them. It is people who are working in these areas that must determine the rights and wrongs of proposals. It is the duty of UTC to see whether it is the prestige of 2 or 3 people, or the logical encoding of the language, that has greater importance.

Only the handful of people in and revolving around the committee in Kerala support the proposed atomic chillu codepoints, whereas the whole of scholarly community, including senior language and computer experts, does not accept it and is held in derision by the community.

It is important to note that, the said committee who produced the report and the initial proposers of atomic chillus, have not given any rationale for it. Reasonings were produced after the fact, by a group of persons who wish to establish their identity in important forums like Unicode. They bring "evidences" and "words" with no attested use or meaning in Malayalam to establish their identity. It is all too easy to do the same in Latin scripts.

It is to be noted that, if hypothetical words are to be used to support a proposal, then it is all too easy also to create new conjunct forms, such as chillu-ന + subjoined-post-base-ല (just as in the case of ന്ന).

It is difficult to say it, but some UTC members view this problem in a non-academic way. The UTC asked for expertise. It has been delivered. The UTC asked for consensus. It has been delivered. Now, the UTC and allied bodies, continues with a silly proposal from the point of view that some government committee has decided so. It is a fact that the said committee's report is totally erroneous, and has not been accepted by the people or by the scholarly community. It is this report that was resubmitted without achieving any consensus in Kerala.

It should be noted that even though the recommendations of the committee were accepted in 2001, and it disallows government purchasing of software not conformant to it, most of the so-called Malayalam software purchased by Government since then are non-conformant with the standard.

Dear UTC members, Rachana is not trying to establish something new, rather we have always upheld the original encoding of Unicode which is efficient and economical for the implementation of Malayalam. In effect, we are saying that the standard formulated by ISCII and accepted by UTC is absolutely correct, and on the other hand, the UTC claims it is wrong. On the one side, repeated pursuit of chillu codepoints, and on the other, logical argument from both IT and linguistics is before the UTC. We request that UTC arrive at a scholarly and logical decision.

Regards,

R. Chitrajakumar,  
N. Gangadharan,  
Rajeev J Sebastian,  
Rachana Aksharavedi